```
  1 CGACAGTCTTTAGTAGGGAAAGGAGACAAGTGCTAGCTACTGCCGCCCAAGTGGAAGGTG   60

 61 GGTGAAATTGCTCACTCTTCACCCCACTGACGCTTTTGCGCACCTGGAAAAGCGGTTCCA  120

121 GTTTGCGCCCGTCGCCGCCTTACAGCCGACAGGAGACCAGCGCTACCCAAGTCACGTGGG  180

181 TTCAGCCTGCAGCTTTCTTGGCCCGAAAGGGAATTATCTATAGAGTAAGTATGCTAATCT  240
  1                                                 M  L  I  L    4

241 TGACTAAGACTGCAGGAGTTTTTTTTAAACCATCAAAAAGGAAAGTTTATGAATTTTTAA  300
  5   T  K  T  A  G  V  F  F  K  P  S  K  R  K  V  Y  E  F  L  R   24

301 GAAGTTTTAATTTTCATCCTGGAACACTATTTCTTCATAAAATAGTATTGGGAATTGAAA  360
 25   S  F  N  F  H  P  G  T  L  F  L  H  K  I  V  L  G  I  E  T   44

361 CTAGTTGTGATGATACAGCAGCTGCTGTGGTGGATGAAACTGGAAATGTGTTGGGAGAAG  420
 45   S  C  D  D  T  A  A  A  V  V  D  E  T  G  N  V  L  G  E  A   64

421 CAATACATTCCCAAACTGAAGTTCATTTAAAAACAGGTGGGATTGTTCCTCCAGCAGCTC  480
 65   I  H  S  Q  T  E  V  H  L  K  T  G  G  I  V  P  P  A  A  Q   84

481 AACAGCTTCACAGAGAAAATATTCAACGAATAGTACAAGAAGCTCTTTCTGCCAGTGGAG  540
 85   Q  L  H  R  E  N  I  Q  R  I  V  Q  E  A  L  S  A  S  G  V  104

541 TCTCTCCAAGTGACCTCTCAGCAATTGCAACTACCATAAAACCAGGACTTGCTTTAAGCC  600
105   S  P  S  D  L  S  A  I  A  T  T  I  K  P  G  L  A  L  S  L  124

601 TGGGAGTGGGCTTATCATTTAGCTTACAGCTGGTAGGACAGTTAAAAAAGCCATTCATTC  660
125   G  V  G  L  S  F  S  L  Q  L  V  G  Q  L  K  K  P  F  I  P  144

661 CCATTCATCATATGGAGGCTCATGCACTTACTATTAGGTTGACCAATAAAGTAGAATTTC  720
145   I  H  H  M  E  A  H  A  L  T  I  R  L  T  N  K  V  E  F  P  164

721 CTTTTTTAGTTCTTTTGATTTCTGGAGGTCACTGTCTGTTGGCATTAGTTCAAGGAGTTT  780
165   F  L  V  L  L  I  S  G  G  H  C  L  L  A  L  V  Q  G  V  S  184

781 CAGATTTTCTGCTTCTTGGAAAGTCTTTGGACATAGCACCAGGTGACATGCTTGACAAGG  840
185   D  F  L  L  L  G  K  S  L  D  I  A  P  G  D  M  L  D  K  V  204

841 TGGCAAGAAGACTTTCTTTAATAAAACATCCAGAGTGCTCCACCATGAGTGGTGGGAAAG  900
205   A  R  R  L  S  L  I  K  H  P  E  C  S  T  M  S  G  G  K  A  224
```

FIG. 1A

```
 901  CCATAGAACATTTGGCCAAACAAGGAAATAGATTTCATTTTGACATCAAACCTCCCTTGC   960
 225    I  E  H  L  A  K  Q  G  N  R  F  H  F  D  I  K  P  P  L  H   244

 961  ATCATGCTAAAAATTGTGATTTTTCTTTTACTGGACTTCAACACGTTACTGATAAAATAA  1020
 245    H  A  K  N  C  D  F  S  F  T  G  L  Q  H  V  T  D  K  I  I   264

1021  TAATGAAAAAGGAAAAAGAGGAAGGTATTGAGAAGGGGCAAATCCTGTCTTCAGCAGCAG  1080
 265    M  K  K  E  K  E  E  G  I  E  K  G  Q  I  L  S  S  A  A  D   284

1081  ACATTGCTGCCACAGTACAGCACACAATGGCATGTCATCTTGTGAAAAGAACACATCGGG  1140
 285    I  A  A  T  V  Q  H  T  M  A  C  H  L  V  K  R  T  H  R  A   304

1141  CTATTCTGTTTTGTAAGCAGAGAGACTTGTTACCTCAAAATAATGCAGTACTGGTTGCAT  1200
 305    I  L  F  C  K  Q  R  D  L  L  P  Q  N  N  A  V  L  V  A  S   324

1201  CTGGTGGTGTCGCAAGTAACTTCTATATCCGCAGAGCTCTGGAAATTTTAACAAACGCAA  1260
 325    G  G  V  A  S  N  F  Y  I  R  R  A  L  E  I  L  T  N  A  T   344

1261  CACAGTGCACTTTGTTGTGTCCTCCTCCCAGACTATGCACTGATAATGGCATTATGATTG  1320
 345    Q  C  T  L  L  C  P  P  P  R  L  C  T  D  N  G  I  M  I  A   364

1321  CATGGAATGGTATTGAAAGACTACGTGCTGGCTTGGGCATTTTACATGACATAGAAGGCA  1380
 365    W  N  G  I  E  R  L  R  A  G  L  G  I  L  H  D  I  E  G  I   384

1381  TCCGCTATGAACCAAAATGTCCTCTTGGAGTAGACATATCAAAAGAAGTTGGAGAAGCTT  1440
 385    R  Y  E  P  K  C  P  L  G  V  D  I  S  K  E  V  G  E  A  S   404

1441  CCATAAAAGTACCACAATTAAAAATGGAGATATGATTTCTGCTGTTCAAAAAAGTCCCTA  1500
 405    I  K  V  P  Q  L  K  M  E  I  *                              415

1501  AAGGGTCTCACTCTCTGACCTCAGCTGGAGTACAGTAGCCAGATCACAACTCACTGCAAC  1560

1561  CCTGACTTCCTGAACTCAAGAAATCCTCCTGCCTTAGCCTCTTGAATAGCCGGGACTACA  1620

1621  GGTGTGCATGTCCATGCCCAGCCAACTTTATTTCTATTTTTTGTAGAGACAGGCTCTTGC  1680

1681  CATGTTGCCCGGGCTGGTCCTGAACTGCTGAATTCAAGTGATCCTCCCACCTTGGCCTCC  1740

1741  AGAAGTGCTGGGATTATGGGTGTGAGCCACCATGCCTAGCCAAAATGTTTCTTAAGGTAT  1800

1801  ACATTTTGGGTCTTAGAAGACTTATACATTTGTAATATTTATTACTAAATATCTCAAAGT  1860

1861  ATTACAATAAATGTTACCATGTGAGCTACTTTGAATCAGGCTTCTTGCACACCAATTTAA  1920
```

FIG. 1B

```
1921  AAATGTTAACTCTTGATATATACACTAGTTATACCACTCATGTCAGTCAATAAATTTTAA  1980

1981  GGTTTAAGTGCAGGCCTTTGTTTACAGAAATCCTAATTTTTTGAAACCATAACTCTGACC  2040

2041  TGACACTAAATTCCTGTAGACATGCTAAGGAAAATCTGCTTAGTATCGAGATCAAGAACT  2100

2101  TCCATTCAAAAAGATTATTCAGTTATGTTATTTGCATATTACCATTGTTAAAAATAAAAA  2160

2161  AATTTTTAAAAGATGAAAAAAAAAAAAAAAAAAAAA  2197
```

FIG. 1C

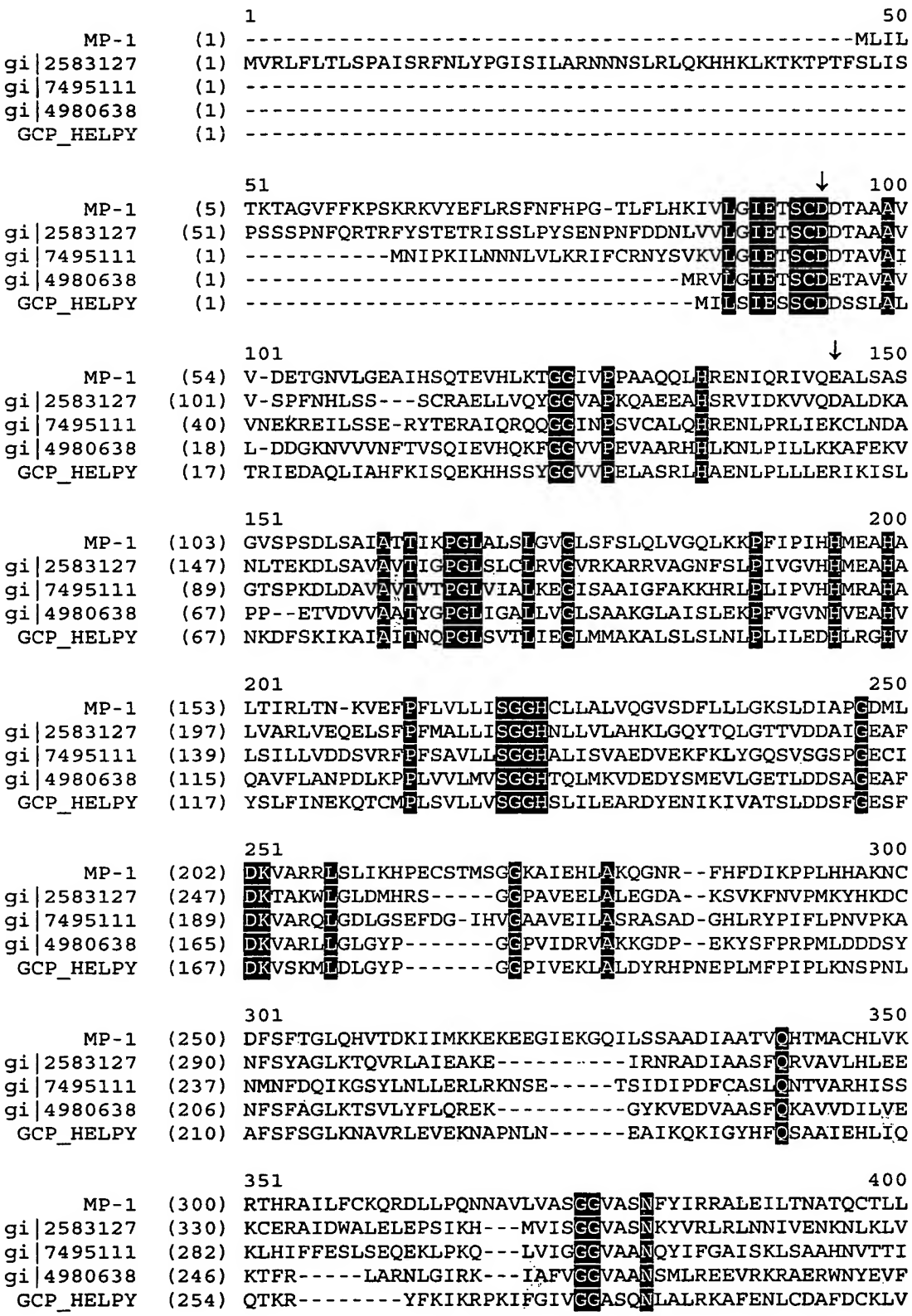
1 50
MP-1 (1) ----------MLIL
gi|2583127 (1) MVRLFLTLSPAISRFNLYPGISILARNNNSLRLQKHHKLKTKTPTFSLIS
gi|7495111 (1) ----------
gi|4980638 (1) ----------
GCP_HELPY (1) ----------
51 100
MP-1 (5) TKTAGVFFKPSKRKVYEFLRSFNFHPG-TLFLHKIVLGIETSCDDTAAAV
gi|2583127 (51) PSSSPNFQRTRFYSTETRISSLPYSENPNFDDNLVVLGIETSCDDTAAAV
gi|7495111 (1) ----------MNIPKILNNNLVLKRIFCRNYSVKVLGIETSCDDTAVAI
gi|4980638 (1) ----------MRVLGIETSCDETAVAV
GCP_HELPY (1) ----------MILSIESSCDDSSLAL
101 150
MP-1 (54) V-DETGNVLGEAIHSQTEVHLKTGGIVPPAAQQLHRENIQRIVQEALSAS
gi|2583127 (101) V-SPFNHLSS---SCRAELLVQYGGVAPKQAEEAHSRVIDKVVQDALDKA
gi|7495111 (40) VNEKREILSSE-RYTERAIQRQQGGINPSVCALQHRENLPRLIEKCLNDA
gi|4980638 (18) L-DDGKNVVVNFTVSQIEVHQKFGGVVPEVAARHHLKNLPILLKKAFEKV
GCP_HELPY (17) TRIEDAQLIAHFKISQEKHHSSYGGVVPELASRLHAENLPLLLERIKISL
151 200
MP-1 (103) GVSPSDLSAIATTIKPGLALSLGVGLSFSLQLVGQLKKPFIPIHHMEAHA
gi|2583127 (147) NLTEKDLSAVAVTIGPGLSLCLRVGVRKARRVAGNFSLPIVGVHHMEAHA
gi|7495111 (89) GTSPKDLDAVAVTVTPGLVIALKEGISAAIGFAKKHRLPLIPVHHMRAHA
gi|4980638 (67) PP--ETVDVVAATYGPGLIGALLVGLSAAKGLAISLEKPFVGVNHVEAHV
GCP_HELPY (67) NKDFSKIKAIAITNQPGLSVTLIEGLMMAKALSLSLNLPLILEDHLRGHV
201 250
MP-1 (153) LTIRLTN-KVEFPFLVLLISGGHCLLALVQGVSDFLLLGKSLDIAPGDML
gi|2583127 (197) LVARLVEQELSFPFMALLISGGHNLLVLAHKLGQYTQLGTTVDDAIGEAF
gi|7495111 (139) LSILLVDDSVRFPFSAVLLSGGHALISVAEDVEKFKLYGQSVSGSPGECI
gi|4980638 (115) QAVFLANPDLKPPLVVLMVSGGHTQLMKVDEDYSMEVLGETLDDSAGEAF
GCP_HELPY (117) YSLFINEKQTCMPLSVLLVSGGHSLILEARDYENIKIVATSLDDSFGESF
251 300
MP-1 (202) DKVARRLSLIKHPECSTMSGGKAIEHLAKQGNR--FHFDIKPPLHHAKNC
gi|2583127 (247) DKTAKWLGLDMHRS-----GGPAVEELALEGDA--KSVKFNVPMKYHKDC
gi|7495111 (189) DKVARQLGDLGSEFDG-IHVGAAVEILASRASAD-GHLRYPIFLPNVPKA
gi|4980638 (165) DKVARLLGLGYP-------GGPVIDRVAKKGDP--EKYSFPRPMLDDDSY
GCP_HELPY (167) DKVSKMLDLGYP-------GGPIVEKLALDYRHPNEPLMFPIPLKNSPNL
301 350
MP-1 (250) DFSFTGLQHVTDKIIMKKEKEEGIEKGQILSSAADIAATVQHTMACHLVK
gi|2583127 (290) NFSYAGLKTQVRLAIEAKE----------IRNRADIAASFQRVAVLHLEE
gi|7495111 (237) NMNFDQIKGSYLNLLERLRKNSE-----TSIDIPDFCASLQNTVARHISS
gi|4980638 (206) NFSFAGLKTSVLYFLQREK----------GYKVEDVAASFQKAVVDILVE
GCP_HELPY (210) AFSFSGLKNAVRLEVEKNAPNLN------EAIKQKIGYHFQSAAIEHLIQ
351 400
MP-1 (300) RTHRAILFCKQRDLLPQNNAVLVASGGVASNFYIRRALEILTNATQCTLL
gi|2583127 (330) KCERAIDWALELEPSIKH---MVISGGVASNKYVRLRLNNIVENKNLKLV
gi|7495111 (282) KLHIFFESLSEQEKLPKQ---LVIGGGVAANQYIFGAISKLSAAHNVTTI
gi|4980638 (246) KTFR-----LARNLGIRK---IAFVGGVAANSMLREEVRKRAERWNYEVF
GCP_HELPY (254) QTKR--------YFKIKRPKIFGIVGGASQNLALRKAFENLCDAFDCKLV

FIG. 2A

```
                    401                                                 450
      MP-1   (350) CPPPRLCTDNGIMIAWNGIERLR-AGLGILHDIEGIRYEPKCPLGVDISK
gi|2583127   (377) CPPPSLCTDNGVMVAWTGLEHFR-VGRYDPPPPATEPEDYVYDLRPRWPL
gi|7495111   (329) KVLLSLCTDNAEMIAYSGLLMLVNRSEAIWWRPNDIPDTIYAHARSDIGT
gi|4980638   (288) FPPLELCTDNALMVAKAGYEKAK-RGMFSPLSLNADPNLNV---------
 GCP_HELPY   (296) LAPLEFCSDNAAMIGRSSLEAYQ-KKRFVPLEKANISPRTLLKSFE----

                    451                                         493
      MP-1   (399) EVGEASIKVPQLKMEI----------------------------
gi|2583127   (426) GEEYAKGRSEARSMRTARIHPSLTSIIRADSLQQQTQT-----
gi|7495111   (379) DASSEIIDTPRRKLVTSTIHGTERIRFRNLDDFKKPKSPKTTE
gi|4980638   (328) ----------------------------------------
 GCP_HELPY   (341) ----------------------------------------
```

FIG. 2B

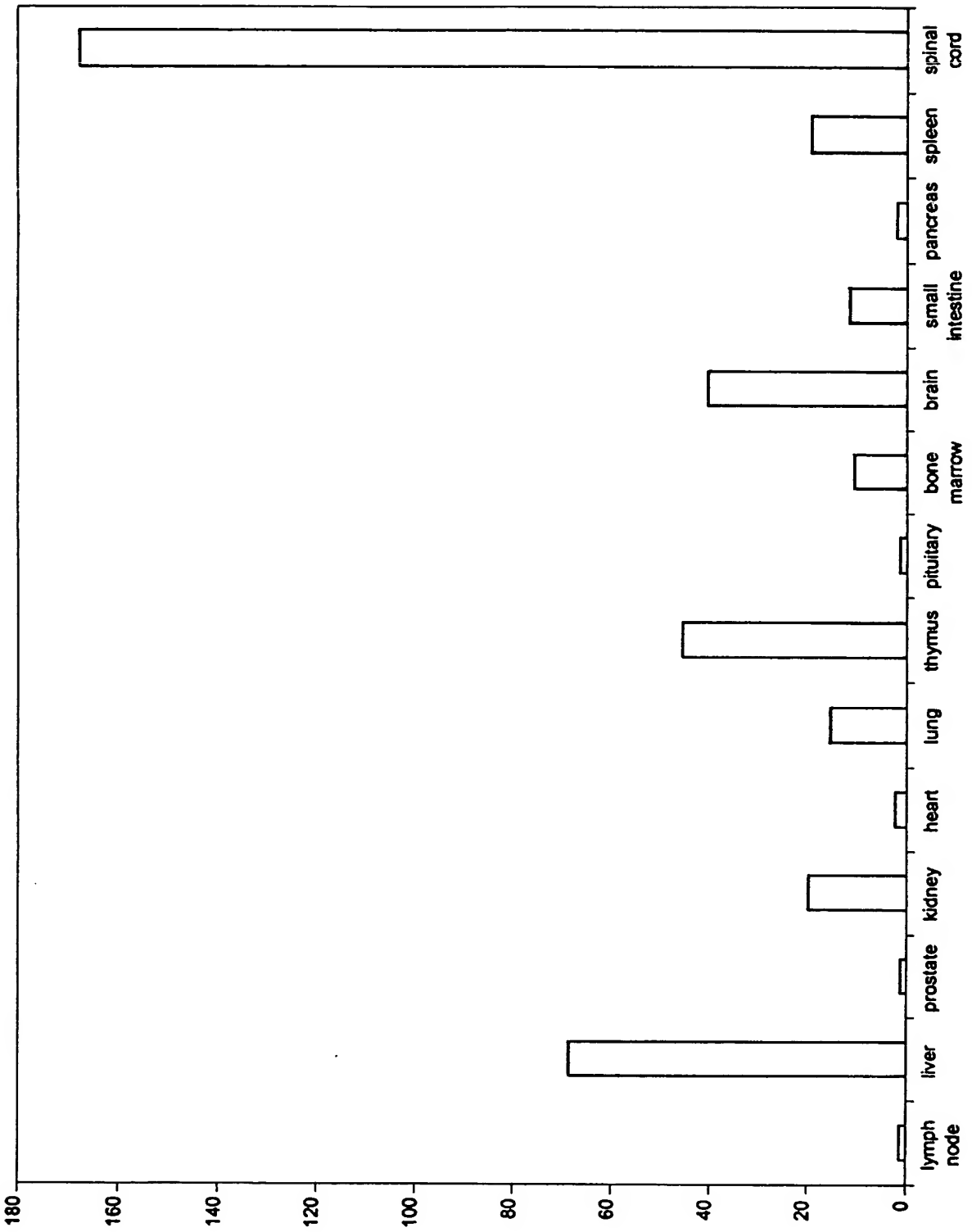
180
160
140
120
100
80
60
40
20
0
lymph node
liver
prostate
kidney
heart
lung
thymus
pituitary
bone marrow
brain
small intestine
pancreas
spleen
spinal cord

FIG. 3

```
                1                                                    *50
MP-1     (1)    MLILTKTAGVFFKPSKRKVYEFLRSFNFHPGTLFLHKIVLGIETSCDDTA
HYPD     (1)    --------------------------------------------------

               51                                                   *100
MP-1    (51)    AAVVDETGNVLGEAIHSQTEVHLKTGGIVPPAAQQLHRENIQRIVQEALS
HYPD     (1)    --------------------------------------------------

              101                                                  * 150
MP-1   (101)    ASGVSPSDLSAIATTIKPGLALSLGVGLSFSLQLVGQLKKPFIPIHHMEA
HYPD     (1)    --------------------------------------------------

              151                                                   200
MP-1   (151)    HALTIRLTNKVEFPFLVLLISGGHCLLALVQGVSDFLLLGKSLDIAPGDM
HYPD     (1)    ------------------------MRILVLGVGNILLTDEAIGVRIVEA

              201                                                   250
MP-1   (201)    LDKVARRLSLIKHPECSTMSGGKAIEHLAKQGNRFHFDIKPPLHHAKNCD
HYPD    (26)    LEQRYILPDYVEILDGGTAG----MELLGDMANRDHLIIADAIVSKKNAP

              251                                                   300
MP-1   (251)    FSFTGLQHVTDKIIMKKEKEEGIEKGQILSSAADIAATVQHTMACHLVKR
HYPD    (72)    GTMMILRDEEVPALFTN----KISPHQLGLADVLSALRFTGEFPKKLTLV

              301                                                   350
MP-1   (301)    THRAILFCKQRDLLPQNNAVLVASGGVASNFYIRRALEILTNATQCTLLC
HYPD   (118)    GVIPESLEPHIGLTPTVEAMIEPALEQVLAALRESGVEAIPRSDS-----

              351                                                   400
MP-1   (351)    PPPRLCTDNGIMIAWNGIERLRAGLGILHDIEGIRYEPKCPLGVDISKEV
HYPD   (163)    --------------------------------------------------

              401        414
MP-1   (401)    GEASIKVPQLKMEI
HYPD   (163)    --------------
```

FIG. 4

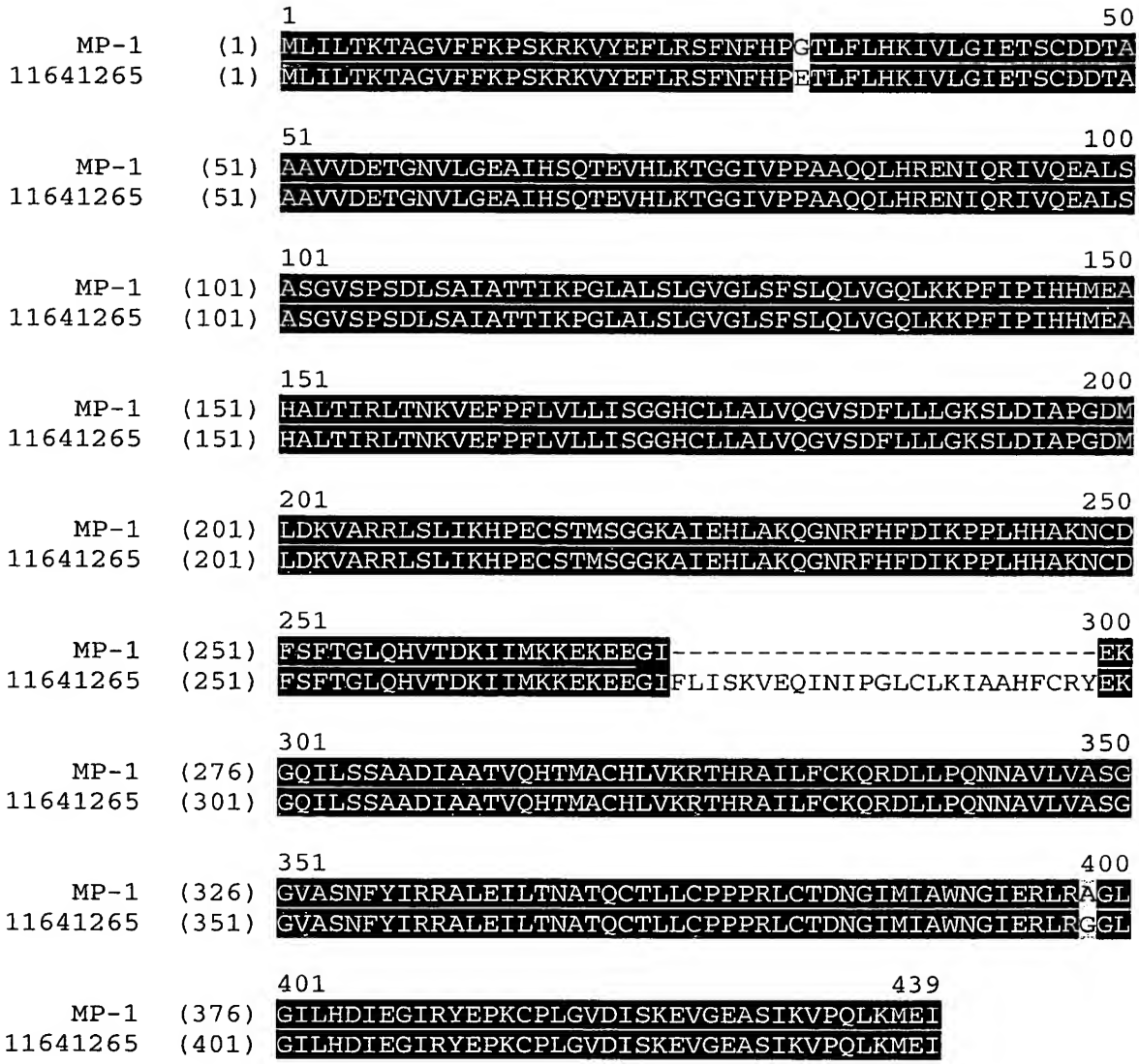
1 50
MP-1 (1) MLILTKTAGVFFKPSKRKVYEFLRSFNFHPGTLFLHKIVLGIETSCDDTA
11641265 (1) MLILTKTAGVFFKPSKRKVYEFLRSFNFHPETLFLHKIVLGIETSCDDTA
51 100
MP-1 (51) AAVVDETGNVLGEAIHSQTEVHLKTGGIVPPAAQQLHRENIQRIVQEALS
11641265 (51) AAVVDETGNVLGEAIHSQTEVHLKTGGIVPPAAQQLHRENIQRIVQEALS
101 150
MP-1 (101) ASGVSPSDLSAIATTIKPGLALSLGVGLSFSLQLVGQLKKPFIPIHHMEA
11641265 (101) ASGVSPSDLSAIATTIKPGLALSLGVGLSFSLQLVGQLKKPFIPIHHMEA
151 200
MP-1 (151) HALTIRLTNKVEFPFLVLLISGGHCLLALVQGVSDFLLLGKSLDIAPGDM
11641265 (151) HALTIRLTNKVEFPFLVLLISGGHCLLALVQGVSDFLLLGKSLDIAPGDM
201 250
MP-1 (201) LDKVARRLSLIKHPECSTMSGGKAIEHLAKQGNRFHFDIKPPLHHAKNCD
11641265 (201) LDKVARRLSLIKHPECSTMSGGKAIEHLAKQGNRFHFDIKPPLHHAKNCD
251 300
MP-1 (251) FSFTGLQHVTDKIIMKKEKEEGI----------EK
11641265 (251) FSFTGLQHVTDKIIMKKEKEEGIFLISKVEQINIPGLCLKIAAHFCRYEK
301 350
MP-1 (276) GQILSSAADIAATVQHTMACHLVKRTHRAILFCKQRDLLPQNNAVLVASG
11641265 (301) GQILSSAADIAATVQHTMACHLVKRTHRAILFCKQRDLLPQNNAVLVASG
351 400
MP-1 (326) GVASNFYIRRALEILTNATQCTLLCPPPRLCTDNGIMIAWNGIERLRAGL
11641265 (351) GVASNFYIRRALEILTNATQCTLLCPPPRLCTDNGIMIAWNGIERLRGGL
401 439
MP-1 (376) GILHDIEGIRYEPKCPLGVDISKEVGEASIKVPQLKMEI
11641265 (401) GILHDIEGIRYEPKCPLGVDISKEVGEASIKVPQLKMEI

FIG. 5

```
mp1                1 MLILTKTAGVFFKPSKRKVYEFLRSFNFHPGTLFLHKIVLGIETSCDDT
                     MLILTKTAGVFFKPSKRKVYEFLRSFNFHPGTLFLHKIVLGIETSCDDT
                     MLILTKTAGVFFKPSKRKVYEFLRSFNFHPGTLFLHKIVLGIETSCDDT
AC013468_6      -9850 acataaagggttactaaagtgttaatatccgactccaagtgagaatgga
                      ttttcaccgtttaccagataattggtatacgctttaatttgtacggaac
                      gacgtgtaatttaaaagattataatttttaaatttaaagatattttta

mp1               50 AAAVVDETGNVLGEAIHSQTEVHL
                     AAAVVDETGNVLGEAIHSQTEVHL
                     AAAVVDETGNVLGEAIHSQTEVHL           K:K[aaa]
AC013468_6      -9703 gggggggagagtgggactcaggctAAGTAAGTA  Intron 1
                      cccttaacgattgactacacatat  <2-----[9629 : 3771]
                      attggtatatggaaaatcatatta

mp1               74     TGGIVPPAAQQLHRENIQRIVQEALSASGVSPSDLSAIATTIKPG
                         TGGIVPPAAQQLHRENIQRIVQEALSASGVSPSDLSAIATTIKPG
                         TGGIVPPAAQQLHRENIQRIVQEALSASGVSPSDLSAIATTIKPG
AC013468_6      -3773 CAGAaggagccggccccagaaccagcggctgaggtcagctgagaaaacg
                      -2> cggttccccaatagaatagttaactccggtccgatcctccctacg
                          atgtttaatagtcaattaaaaaatttctactatccaatatcaaaa

mp1              120 LALSLGVGLSFSLQLVGQLKKPFIPIHHMEAHALTIRLTNKVEFPFLVL
                     LALSLGVGLSFSLQLVGQLKKPFIPIHHMEAHALTIRLTNKVEFPFLVL
                     LALSLGVGLSFSLQLVGQLKKPFIPIHHMEAHALTIRLTNKVEFPFLVL
AC013468_6      -3634 cgtacgggtttatccggctaactacaccaggcgcaaataaaggtcttgc
                      tctgtgtgtctgtattgataacttctaatacactctgtcaatatctttt
                      ttacgagcaatcaggaagaagactctttggttatttggctaaatttatt

mp1              169 LISGGHCLLALVQGVSDFLLLGKSLDIAPGDMLDK
                     LISGGHCLLALVQGVSDFLLLGKSLDIAPGDMLDK
                     LISGGHCLLALVQGVSDFLLLGKSLDIAPGDMLDK
AC013468_6      -3487 tatggctctgtgcggtgtcccgattgagcggacga
                      ttcggagttcttagtcattttgactatccgattaa
                      gttatctggaataatattgttagtgcaaatcgtcg

mp1              204                          VARRLSLIKHPECSTMSGGKAIEHLA
                                              VARRLSLIKHPECSTMSGGKAIEHLA
                                              VARRLSLIKHPECSTMSGGKAIEHLA
AC013468_6      -3382 GTAATTA  Intron 2   TAGgaacttaaccgttaaaggagagctg
                      <0-----[3382 : 2556]-0>tcggtcttaacagcctgggactaatc
                                              gaaattaaatagcccgttgacaatgc

mp1              230 KQGNRFHFDIKPPLHHAKNCDFSFTGLQHVTDKIIMKKEKEE
                     KQGNRFHFDIKPPLHHAKNCDFSFTGLQHVTDKIIMKKEKEE
                     KQGNRFHFDIKPPLHHAKNCDFSFTGLQHVTDKIIMKKEKEE
AC013468_6      -2477 acgaatctgaacctccgaatgtttagcccgagaaaaaagagg
                      aagagtatatacctaacaagatctcgtaatcaatttaaaaaa
                      aaatatttccatcgtttatttttttatactttaaagagaaga
```

FIG. 6A

```
mp1              272                                IEKGQILSSAADIAATVQHTMAC
                                                    IEKGQILSSAADIAATVQHTMAC
                                 G:G[ggt]           IEKGQILSSAADIAATVQHTMAC
AC013468_6      -2351 GGTATATT  Intron 3   CAGGTagagcacttgggaggagccaagt
                       <1-----[2350 : 2275]-1>  taagattccccatccctaactcg
                                                tgggacgtaaacttcaagcagat

mp1              296 HLVKRTHRAILFCKQRDLLPQNNAVL
                     HLVKRTHRAILFCKQRDLLPQNNAVL
                     HLVKRTHRAILFCKQRDLLPQNNAVL
AC013468_6      -2203 ccgaaaccgacttacagttccaaggcGTAAGTT  Intron 4   CAG
                      attagcagctttgaagattcaaactt<0-----[2125 : 1190]-0>
                      ttgaaatgttgttggacgatattaag

mp1              322 VASGGVASNFYIRRALEILTNATQCTLLCPPPRLCTDNGIMIA
                     VASGGVASNFYIRRALEILTNATQCTLLCPPPRLCTDNGIMIA
                     VASGGVASNFYIRRALEILTNATQCTLLCPPPRLCTDNGIMIA
AC013468_6      -1189 ggtggggaattacagcgataagactatttcccactagagaaag
                      tccggtcgatatggctattcaccagcttgcccgtgcaagtttc
                      tatttcatcctccatgataataagctggtttcaactttctgta

mp1              365                              NGIERLRAGLGILHDIEGIRYEP
                                                  NGIERLRAGLGILHDIEGIRYEP
                               W:W[tgg]           NGIERLRAGLGILHDIEGIRYEP
AC013468_6      -1060 TGGTAAGCC  Intron 5   TAGGagagaccggtgatcgaggactgc
                        <2-----[1058 :  935]-2> agtagtgcgtgttaatagtgaac
                                                tttaaattcgctatcaaccctaa

mp1              389 K
                     K
                     K
AC013468_6       -864 a
                      a
                      g
```

FIG. 6B

```
11641265          1 MLILTKTAGVFFKPSKRKVYEFLRSFNFHPETLFLHKIVLGIETSCDDT
                    MLILTKTAGVFFKPSKRKVYEFLRSFNFHP TLFLHKIVLGIETSCDDT
                    MLILTKTAGVFFKPSKRKVYEFLRSFNFHPGTLFLHKIVLGIETSCDDT
AC013468_6     -9850 acataaagggttactaaagtgttaatatccgactccaagtgagaatgga
                    ttttcaccgtttaccagataattggtatacgctttaatttgtacggaac
                    gacgtgtaatttaaaagattataatttttaaatttaaagatattttta

11641265         50 AAAVVDETGNVLGEAIHSQTEVHL
                    AAAVVDETGNVLGEAIHSQTEVHL
                    AAAVVDETGNVLGEAIHSQTEVHL           K:K[aaa]
AC013468_6     -9703 gggggggagagtgggactcaggctAAGTAAGTA  Intron 1
                    cccttaacgattgactacacatat  <2-----[9629 · 3771]
                    attggtatatggaaaatcatatta

11641265         74     TGGIVPPAAQQLHRENIQRIVQEALSASGVSPSDLSAIATTIKPG
                        TGGIVPPAAQQLHRENIQRIVQEALSASGVSPSDLSAIATTIKPG
                        TGGIVPPAAQQLHRENIQRIVQEALSASGVSPSDLSAIATTIKPG
AC013468_6     -3773 CAGAaggagccggccccagaaccagcggctgaggtcagctgagaaaacg
                    -2> cggttccccaatagaatagttaactccggtccgatcctccctacg
                        atgtttaatagtcaattaaaaaatttctactatccaatatcaaaa

11641265        120 LALSLGVGLSFSLQLVGQLKKPFIPIHHMEAHALTIRLTNKVEFPFLVL
                    LALSLGVGLSFSLQLVGQLKKPFIPIHHMEAHALTIRLTNKVEFPFLVL
                    LALSLGVGLSFSLQLVGQLKKPFIPIHHMEAHALTIRLTNKVEFPFLVL
AC013468_6     -3634 cgtacgggtttatccggctaactacaccaggcgcaaataaaggtcttgc
                    tctgtgtgtctgtattgataacttctaatacactctgtcaatatctttt
                    ttacgagcaatcaggaagaagactctttggttatttggctaaatttatt

11641265        169 LISGGHCLLALVQGVSDFLLLGKSLDIAPGDMLDK
                    LISGGHCLLALVQGVSDFLLLGKSLDIAPGDMLDK
                    LISGGHCLLALVQGVSDFLLLGKSLDIAPGDMLDK
AC013468_6     -3487 tatggctctgtgcggtgtcccgattgagcggacga
                    ttcggagttcttagtcattttgactatccgattaa
                    gttatctggaataatattgttagtgcaaatcgtcg

11641265        204                          VARRLSLIKHPECSTMSGGKAIEHLA
                                             VARRLSLIKHPECSTMSGGKAIEHLA
                                             VARRLSLIKHPECSTMSGGKAIEHLA
AC013468_6     -3382 GTAATTA  Intron 2   TAGgaacttaaccgttaaaggagagctg
                    <0-----[3382 : 2556]-0>tcggtcttaacagcctgggactaatc
                                             gaaattaaatagcccgttgacaatgc

11641265        230 KQGNRFHFDIKPPLHHAKNCDFSFTGLQHVTDKIIMKKEKEEGIFLISK
                    KQGNRFHFDIKPPLHHAKNCDFSFTGLQHVTDKIIMKKEKEEGIFLISK
                    KQGNRFHFDIKPPLHHAKNCDFSFTGLQHVTDKIIMKKEKEEGIFLISK
AC013468_6     -2477 acgaatctgaacctccgaatgtttagcccgagaaaaaagagggatcaaa
                    aagagtatatacctaacaagatctcgtaatcaatttaaaaaagttttga
                    aaatatttccatcgtttatttttttatactttaaagagaagatatatta

FRAME SHIFT HERE
```

FIG. 7A

```
11641265          279 VEQINIPGLCLKIAAHFCRYEKGQILSSAADIAATVQHTMACHLVKRTH
                      VEQINIPGLCLKIAAHFCRY KGQILSSAADIAATVQHTMACHLVKRTH
                      VEQINIPGLCLKIAAHFCRY!KGQILSSAADIAATVQHTMACHLVKRTH
AC013468_6       -2330 ggcaaacgttcaaggcttat4agcacttgggaggagccaagtccgaaac
                      taatatcgtgtatccatgga agattccccatccctaactcgattagca
                      tagatttagcaaatttccgt ggacgtaaacttcaagcagatttgaaat

11641265          328 RAILFCKQRDLLPQNNAVL                          VASGGVA
                      RAILFCKQRDLLPQNNAVL                          VASGGVA
                      RAILFCKQRDLLPQNNAVL                          VASGGVA
AC013468_6       -2182 cgacttacagttccaaggcGTAAGTT  Intron 3   CAGggtgggg
                      gctttgaagattcaaactt<0-----[2125 : 1190]-0>tccggtc
                      gttgttggacgatattaag                          tatttca

11641265          354 SNFYIRRALEILTNATQCTLLCPPPRLCTDNGIMIA
                      SNFYIRRALEILTNATQCTLLCPPPRLCTDNGIMIA
                      SNFYIRRALEILTNATQCTLLCPPPRLCTDNGIMIA
AC013468_6       -1168 aattacagcgataagactatttcccactagagaaag
                      gatatggctattcaccagcttgcccgtgcaagtttc
                      tcctccatgataataagctggtttcaactttctgta

11641265          390                          NGIERLRGGLGILHDIEGIRYEP
                                               NGIERLR GLGILHDIEGIRYEP
                                 W:W[tgg]      NGIERLRAGLGILHDIEGIRYEP
AC013468_6       -1060 TGGTAAGCC  Intron 4   TAGGagagaccggtgatcgaggactgc
                        <2-----[1058 :  935]-2> agtagtgcgtgttaatagtgaac
                                               tttaaattcgctatcaaccctaa

11641265          414 K
                      K
                      K
AC013468_6        -864 a
                      a
                      g
```

FIG. 7B

Protein	Genbank ID	Identities	Similarities
Arabidopsis O-sialoglycoprotein endopeptidase protein	gi|2583127	30.8%	36%
Caenorhabditis glycoproteinase family member protein	gi|7495111	26.6%	41%
Thermotoga secreted metalloendopeptidase Gcp protein	gi|4980638	27.2%	39%
Helicobacter O-sialoglycoprotein endopeptidase GCP_HELPY protein	gi|2499846	22.5%	34%

FIG. 8

D0073 CNT

FIG. 9

```
                    1                                                  50
       MP-1     (1) CGACAGTCTTTAGTAGGGAAAGGAGACAAGTGCTAGCTACTGCCGCCCAA
gi|11641264     (1) --------------------------------------------------

                   51                                                 100
       MP-1    (51) GTGGAAGGTGGGTGAAATTGCTCACTCTTCACCCCACTGACGCTTTTGCG
gi|11641264     (1) --------------------------------------------------

                  101                                                 150
       MP-1   (101) CACCTGGAAAAGCGGTTCCAGTTTGCGCCCGTCGCCGCCTTACAGCCGAC
gi|11641264     (1) --------------------------------------------------

                  151                                                 200
       MP-1   (151) AGGAGACCAGCGCTACCCAAGTCACGTGGGTTCAGCCTGCAGCTTTCTTG
gi|11641264     (1) --------------------------------------------------

                  201                                                 250
       MP-1   (201) GCCCGAAAGGGAATTATCTATAGAGTAAGTATGCTAATCTTGACTAAGAC
gi|11641264     (1) -------CAGGAATTATCTATAGAGTAAGTATGCTAATCTTGACTAAGAC

                  251                                                 300
       MP-1   (251) TGCAGGAGTTTTTTTTAAACCATCAAAAAGGAAAGTTTATGAATTTTTAA
gi|11641264    (44) TGCAGGAGTTTTTTTTAAACCATCAAAAAGGAAAGTTTATGAATTTTTAA

                  301                                                 350
       MP-1   (301) GAAGTTTTAATTTTCATCCTGGAACACTATTTCTTCATAAAATAGTATTG
gi|11641264    (94) GAAGTTTTAATTTTCATCCTGAAACACTATTTCTTCATAAAATAGTATTG

                  351                                                 400
       MP-1   (351) GGAATTGAAACTAGTTGTGATGATACAGCAGCTGCTGTGGTGGATGAAAC
gi|11641264   (144) GGAATTGAAACTAGTTGTGATGATACAGCAGCTGCTGTGGTGGATGAAAC

                  401                                                 450
       MP-1   (401) TGGAAATGTGTTGGGAGAAGCAATACATTCCCAAACTGAAGTTCATTTAA
gi|11641264   (194) TGGAAATGTGTTGGGAGAAGCAATACATTCCCAAACTGAAGTTCATTTAA

                  451                                                 500
       MP-1   (451) AAACAGGTGGGATTGTTCCTCCAGCAGCTCAACAGCTTCACAGAGAAAAT
gi|11641264   (244) AAACAGGTGGGATTGTTCCTCCAGCAGCTCAACAGCTTCACAGAGAAAAT

                  501                                                 550
       MP-1   (501) ATTCAACGAATAGTACAAGAAGCTCTTTCTGCCAGTGGAGTCTCTCCAAG
gi|11641264   (294) ATTCAACGAATAGTACAAGAAGCTCTTTCTGCCAGTGGAGTCTCTCCAAG

                  551                                                 600
       MP-1   (551) TGACCTCTCAGCAATTGCAACTACCATAAAACCAGGACTTGCTTTAAGCC
gi|11641264   (344) TGACCTCTCAGCAATTGCAACTACCATAAAACCAGGACTTGCTTTAAGCC

                  601                                                 650
       MP-1   (601) TGGGAGTGGGCTTATCATTTAGCTTACAGCTGGTAGGACAGTTAAAAAAG
gi|11641264   (394) TGGGAGTGGGCTTATCATTTAGCTTACAGCTGGTAGGACAGTTAAAAAAG

                  651                                                 700
       MP-1   (651) CCATTCATTCCCATTCATCATATGGAGGCTCATGCACTTACTATTAGGTT
gi|11641264   (444) CCATTCATTCCCATTCATCATATGGAGGCTCATGCACTTACTATTAGGTT

                  701                                                 750
       MP-1   (701) GACCAATAAAGTAGAATTTCCTTTTTTAGTTCTTTTGATTTCTGGAGGTC
gi|11641264   (494) GACCAATAAAGTAGAATTTCCTTTTTTAGTTCTTTTGATTTCTGGAGGTC

                  751                                                 800
       MP-1   (751) ACTGTCTGTTGGCATTAGTTCAAGGAGTTTCAGATTTTCTGCTTCTTGGA
gi|11641264   (544) ACTGTCTGTTGGCATTAGTTCAAGGAGTTTCAGATTTTCTGCTTCTTGGA

                  801                                                 850
       MP-1   (801) AAGTCTTTGGACATAGCACCAGGTGACATGCTTGACAAGGTGGCAAGAAG
gi|11641264   (594) AAGTCTTTGGACATAGCACCAGGTGACATGCTTGACAAGGTGGCAAGAAG
```

FIG. 10A

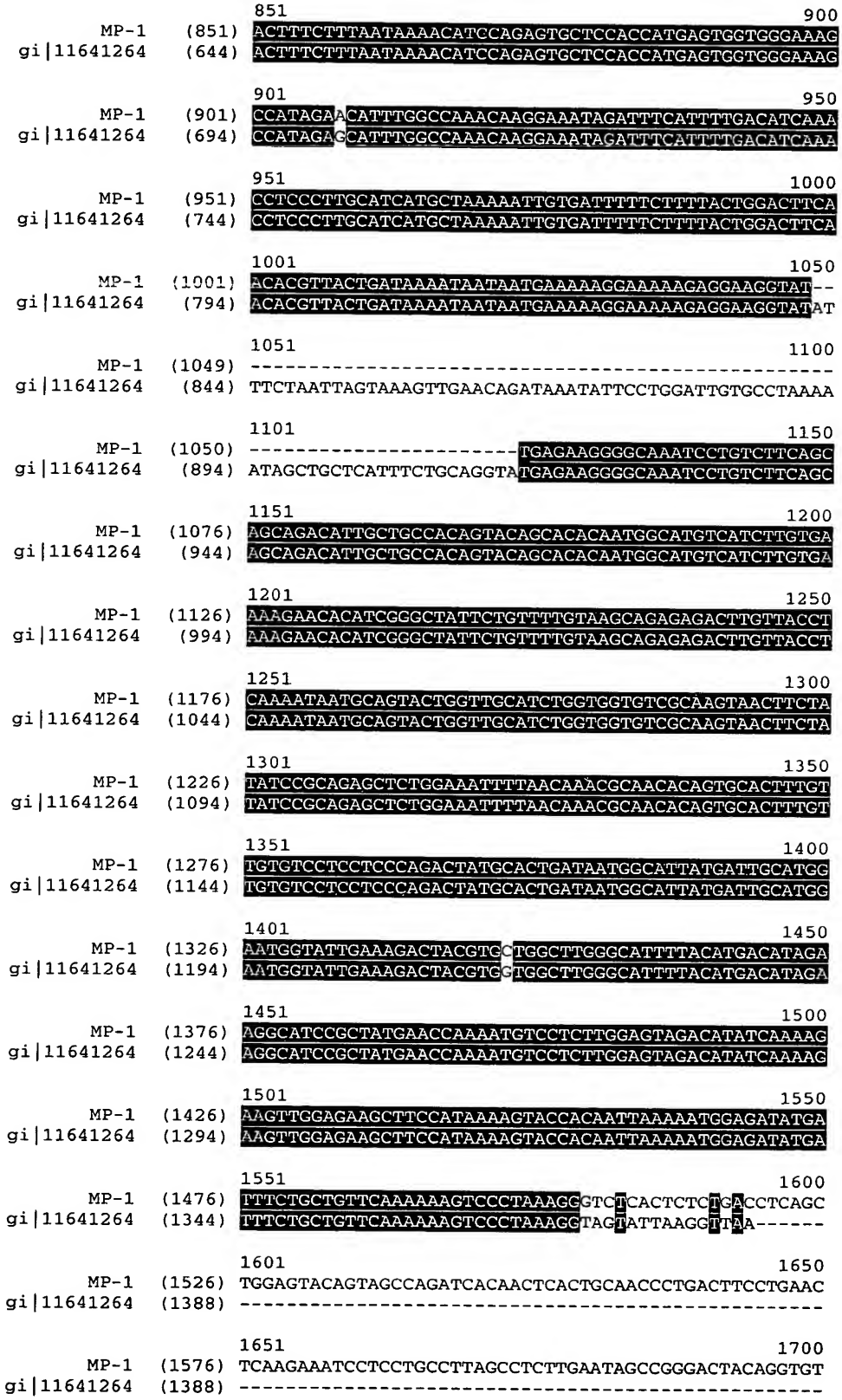
851 900
MP-1 (851) ACTTTCTTTAATAAAACATCCAGAGTGCTCCACCATGAGTGGTGGGAAAG
gi|11641264 (644) ACTTTCTTTAATAAAACATCCAGAGTGCTCCACCATGAGTGGTGGGAAAG
901 950
MP-1 (901) CCATAGAACATTTGGCCAAACAAGGAAATAGATTTCATTTTGACATCAAA
gi|11641264 (694) CCATAGAGCATTTGGCCAAACAAGGAAATAGATTTCATTTTGACATCAAA
951 1000
MP-1 (951) CCTCCCTTGCATCATGCTAAAAATTGTGATTTTTCTTTTACTGGACTTCA
gi|11641264 (744) CCTCCCTTGCATCATGCTAAAAATTGTGATTTTTCTTTTACTGGACTTCA
1001 1050
MP-1 (1001) ACACGTTACTGATAAAATAATAATGAAAAAGGAAAAAGAGGAAGGTAT--
gi|11641264 (794) ACACGTTACTGATAAAATAATAATGAAAAAGGAAAAAGAGGAAGGTATAT
1051 1100
MP-1 (1049) ----------
gi|11641264 (844) TTCTAATTAGTAAAGTTGAACAGATAAATATTCCTGGATTGTGCCTAAAA
1101 1150
MP-1 (1050) ----------TGAGAAGGGGCAAATCCTGTCTTCAGC
gi|11641264 (894) ATAGCTGCTCATTTCTGCAGGTATGAGAAGGGGCAAATCCTGTCTTCAGC
1151 1200
MP-1 (1076) AGCAGACATTGCTGCCACAGTACAGCACACAATGGCATGTCATCTTGTGA
gi|11641264 (944) AGCAGACATTGCTGCCACAGTACAGCACACAATGGCATGTCATCTTGTGA
1201 1250
MP-1 (1126) AAAGAACACATCGGGCTATTCTGTTTTGTAAGCAGAGAGACTTGTTACCT
gi|11641264 (994) AAAGAACACATCGGGCTATTCTGTTTTGTAAGCAGAGAGACTTGTTACCT
1251 1300
MP-1 (1176) CAAAATAATGCAGTACTGGTTGCATCTGGTGGTGTCGCAAGTAACTTCTA
gi|11641264 (1044) CAAAATAATGCAGTACTGGTTGCATCTGGTGGTGTCGCAAGTAACTTCTA
1301 1350
MP-1 (1226) TATCCGCAGAGCTCTGGAAATTTTAACAAACGCAACACAGTGCACTTTGT
gi|11641264 (1094) TATCCGCAGAGCTCTGGAAATTTTAACAAACGCAACACAGTGCACTTTGT
1351 1400
MP-1 (1276) TGTGTCCTCCTCCCAGACTATGCACTGATAATGGCATTATGATTGCATGG
gi|11641264 (1144) TGTGTCCTCCTCCCAGACTATGCACTGATAATGGCATTATGATTGCATGG
1401 1450
MP-1 (1326) AATGGTATTGAAAGACTACGTGCTGGCTTGGGCATTTTACATGACATAGA
gi|11641264 (1194) AATGGTATTGAAAGACTACGTGGTGGCTTGGGCATTTTACATGACATAGA
1451 1500
MP-1 (1376) AGGCATCCGCTATGAACCAAAATGTCCTCTTGGAGTAGACATATCAAAAG
gi|11641264 (1244) AGGCATCCGCTATGAACCAAAATGTCCTCTTGGAGTAGACATATCAAAAG
1501 1550
MP-1 (1426) AAGTTGGAGAAGCTTCCATAAAAGTACCACAATTAAAAATGGAGATATGA
gi|11641264 (1294) AAGTTGGAGAAGCTTCCATAAAAGTACCACAATTAAAAATGGAGATATGA
1551 1600
MP-1 (1476) TTTCTGCTGTTCAAAAAAGTCCCTAAAGGGTCTCACTCTCTGACCTCAGC
gi|11641264 (1344) TTTCTGCTGTTCAAAAAAGTCCCTAAAGGTAGTATTAAGGTTAA------
1601 1650
MP-1 (1526) TGGAGTACAGTAGCCAGATCACAACTCACTGCAACCCTGACTTCCTGAAC
gi|11641264 (1388) ----------
1651 1700
MP-1 (1576) TCAAGAAATCCTCCTGCCTTAGCCTCTTGAATAGCCGGGACTACAGGTGT
gi|11641264 (1388) ----------

FIG. 10B

```
                    1701                                                1750
       MP-1  (1626) GCATGTCCATGCCCAGCCAACTTTATTTCTATTTTTTGTAGAGACAGGCT
gi|11641264  (1388) ----------------------------------------------------

                    1751                                                1800
       MP-1  (1676) CTTGCCATGTTGCCCGGGCTGGTCCTGAACTGCTGAATTCAAGTGATCCT
gi|11641264  (1388) ----------------------------------------------------

                    1801                                                1850
       MP-1  (1726) CCCACCTTGGCCTCCAGAAGTGCTGGGATTATGGGTGTGAGCCACCATGC
gi|11641264  (1388) ----------------------------------------------------

                    1851                                                1900
       MP-1  (1776) CTAGCCAAAATGTTTCTTAAGGTATACATTTTGGGTCTTAGAAGACTTAT
gi|11641264  (1388) ----------------------------------------------------

                    1901                                                1950
       MP-1  (1826) ACATTTGTAATATTTATTACTAAATATCTCAAAGTATTACAATAAATGTT
gi|11641264  (1388) ----------------------------------------------------

                    1951                                                2000
       MP-1  (1876) ACCATGTGAGCTACTTTGAATCAGGCTTCTTGCACACCAATTTAAAAATG
gi|11641264  (1388) ----------------------------------------------------

                    2001                                                2050
       MP-1  (1926) TTAACTCTTGATATATACACTAGTTATACCACTCATGTCAGTCAATAAAT
gi|11641264  (1388) ----------------------------------------------------

                    2051                                                2100
       MP-1  (1976) TTTAAGGTTTAAGTGCAGGCCTTTGTTTACAGAAATCCTAATTTTTTGAA
gi|11641264  (1388) ----------------------------------------------------

                    2101                                                2150
       MP-1  (2026) ACCATAACTCTGACCTGACACTAAATTCCTGTAGACATGCTAAGGAAAAT
gi|11641264  (1388) ----------------------------------------------------

                    2151                                                2200
       MP-1  (2076) CTGCTTAGTATCGAGATCAAGAACTTCCATTCAAAAAGATTATTCAGTTA
gi|11641264  (1388) ----------------------------------------------------

                    2201                                                2250
       MP-1  (2126) TGTTATTTGCATATTACCATTGTTAAAAATAAAAAAATTTTTAAAAGATG
gi|11641264  (1388) ----------------------------------------------------

                    2251              2272
       MP-1  (2176) AAAAAAAAAAAAAAAAAAAAA
gi|11641264  (1388) ---------------------
```

FIG. 10C

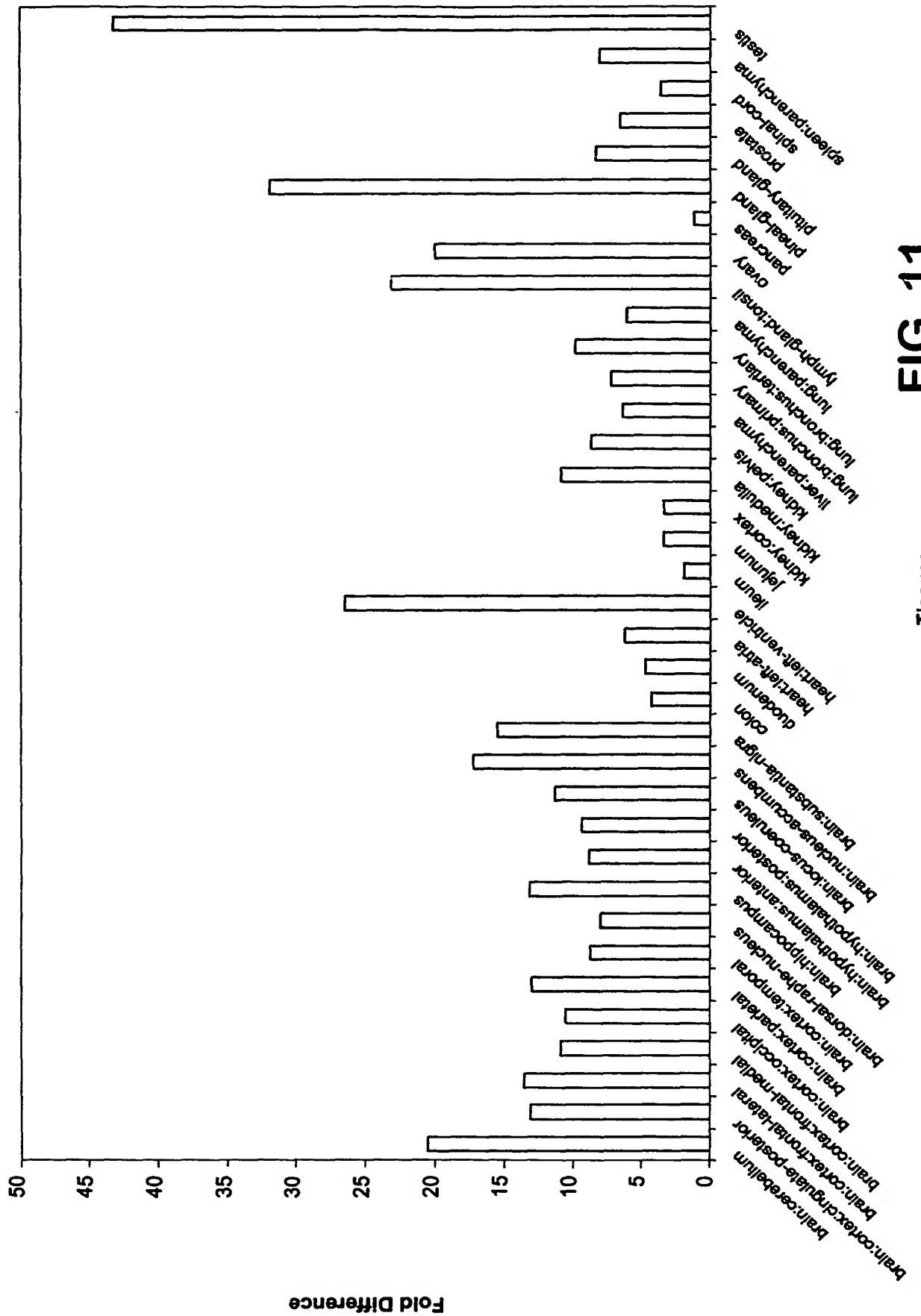
Fold Difference
0
5
10
15
20
25
30
35
40
45
50
brain:cerebellum
brain:cortex:cingulate-posterior
brain:cortex:frontal-lateral
brain:cortex:frontal-medial
brain:cortex:occipital
brain:cortex:parietal
brain:cortex:temporal
brain:dorsal-raphe-nucleus
brain:hippocampus
brain:hypothalamus:anterior
brain:hypothalamus:posterior
brain:locus-coeruleus
brain:nucleus-accumbens
brain:substantia-nigra
colon
duodenum
heart:left-atria
heart:left-ventricle
ileum
jejunum
kidney:cortex
kidney:medulla
liver:pericentral
lung:bronchus:primary
lung:bronchus:tertiary
lymph-gland:tonsil
ovary
pancreas
pineal-gland
pituitary-gland
prostate
spinal-cord
spleen:parenchyma
testis
Tissues

FIG. 11